

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: August 28, 2003, 18:27:17 ; Search time 31.3182 Seconds  
(without alignments)  
107.116 Million cell updates/sec

Title: US-09-743-225-10  
Perfect score: 66  
Sequence: 1 CATLRVYKGGXA 13  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	42	63.6	605	Q9U189	Q9U189 leishmania
2	41	62.1	569	Q8XVH4	Q8XVH4 raietonia s
3	40	60.6	276	Q8D4T7	Q8D4T7 vibrio vuln
4	40	60.6	381	Q8DES0	Q8DES0 vibrio vuln
5	40	60.6	384	Q8Z8V6	Q8Z8V6 salmonella
6	40	60.6	607	Q9ZUR7	Q9ZUR7 rhizobium m
7	40	60.6	619	Q9KKS7	Q9KKS7 vibrio chol
8	40	60.6	932	Q9LXL3	Q9LXL3 arabisidopsis
9	40	60.6	937	Q8LGU3	Q8LGU3 arabisidopsis
10	40	60.6	937	Q8L5J2	Q8L5J2 arabisidopsis
11	40	60.6	938	Q8LNN2	Q8LNN2 arabisidopsis
12	39	59.1	173	Q8IU03	Q8IU03 caenorhabdi
13	39	59.1	241	Q8MXR7	Q8MXR7 caenorhabdi
14	39	59.1	285	Q98EM3	Q98EM3 rhizobium l
15	39	59.1	456	Q8T0Y4	Q8T0Y4 apis mellif
16	39	59.1	631	Q8U8E6	Q8U8E6 agrobacteri

17	38	57.6	147	10	Q9AWT7	Q9awt7 oryza sativ
18	38	57.6	333	10	Q9SYM0	Q9sym0 arabidopsis
19	38	57.6	333	10	Q8LBN5	Q8lbn5 arabidopsis
20	38	57.6	394	17	Q9Y9A1	Q9y9a1 aeropyrum p
21	37	56.1	165	3	Q03083	Q03083 saccharomyc
22	37	56.1	231	16	Q8YT58	Q8yt58 anabaena sp
23	37	56.1	271	16	Q8DUY4	Q8duy4 streptococc
24	37	56.1	273	16	Q8UKN5	Q8ukn5 agrobacteri
25	37	56.1	273	17	Q8TS14	Q8ts14 methanosarc
26	37	56.1	406	10	Q9FN06	Q9fn06 arabidopsis
27	37	56.1	498	10	Q943J5	Q943j5 oryza sativ
28	37	56.1	746	13	Q8JG43	Q8jg43 fuca rubrip
29	37	56.1	822	11	Q8BJ16	Q8bj16 mus musculu
30	37	56.1	835	16	P74210	P74210 synechocyst
31	37	56.1	1104	4	Q92922	Q92922 homo sapien
32	37	56.1	1227	16	Q97K41	Q97k41 clostridium
33	37	56.1	1293	10	Q8H722	Q8h722 oryza sativ
34	36.5	55.3	108	2	Q8KTT0	Q8ktt0 candidatus
35	36.5	55.3	344	5	Q964N0	Q964n0 drosophila
36	36	54.5	73	5	Q9BLJ3	Q9blj3 ornithodora
37	36	54.5	73	5	Q8MY08	Q8my08 ornithodora
38	36	54.5	170	16	Q25227	Q25227 helicobacte
39	36	54.5	175	16	Q9ZL21	Q9z121 helicobacte
40	36	54.5	236	16	Q8UGS4	Q8ugs4 agrobacteri
41	36	54.5	286	16	Q9KEF2	Q9kef2 bacillus ha
42	36	54.5	304	6	Q9N120	Q9n120 oryctolagus
43	36	54.5	523	17	Q97VY8	Q97vy8 sulfobolus
44	36	54.5	533	5	O17592	O17592 caenorhabdi
45	36	54.5	549	2	O31002	O31002 vibrio angu

ALIGNMENTS

RESULT 1

Q9U189 PRELIMINARY; PRT; 605 AA.  
AC Q9U189;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2003 (TREMBLrel. 23, Last annotation update)  
DE Zeta tubulin (Gamma tubulin-related protein).  
GN L3238.01.  
OS Leishmania major.  
OC Eukaryota; Euzoenzoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
OX NCBI\_TaxID=5664;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Friedlin;  
RA Wedler H., Hilbert H., Dueterhoeft A., Ivens A.C., Murphy L.,  
RA Quail M., Rajandream M.A., Barrell B.G.;  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Friedlin;  
RX MEDLINE=98146435; PubMed=9477341;  
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
RA Smith D.F.;  
RL Genome Res. 8:135-145(1998).  
RL EMBL: AL133468; CAB63128.1; -;  
DR InterPro: IPR000896; ER\_target.  
DR InterPro: IPR000217; Tubulin.  
DR InterPro: IPR003008; Tubulin\_FtsZ.  
DR InterPro: IPR004058; Zeta\_tubulin.  
DR Pfam: PF00091; tubulin; 1.  
DR PRINTS: PR01161; TUBULIN.  
DR PRINTS: PR01520; ZETATUBULIN.  
DR PROSITE: PS00014; ER\_TARGET; 1.  
KW GTP-binding.  
SQ SEQUENCE 605 AA; 65700 MW; 49944A374F5D8B82 CRC64;

Query Match 63.6%; Score 42; DB 5; Length 605;

Best Local Similarity 54.5%; Pred. No. 23;  
Matches 6; Conservative 3; Mismatches 0; Gaps 0;

QY 1 CATLRVYKGGG 11  
| : : : : :  
Db 383 CTALRIHEGG 393

## RESULT 2

Q8XVH4 PRELIMINARY; PRT; 569 AA.  
ID Q8XVH4  
AC Q8XVH4  
DT 01-MAR-2002 (TRENBLrel. 20, Created)  
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
DE Probable long-chain fatty-acid--CoA ligase protein (EC 6.2.1.3).  
GN FAD1 OR RSC2857 OR RSC0248.  
OS Ralstonia solanacearum (Pseudomonas solanacearum).  
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
OC Ralstoniaceae; Ralstonia.  
OX NCBI\_TaxID=305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-GM11000;  
RX MEDLINE=21681879; PubMed=11823852;  
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
Ariat M., Billault A., Brotier P., Camus J.C., Catolico L.,  
Chandler C., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,  
Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,  
Weissenbach J., Boucher C.A.;  
RT "Genome sequence of the plant pathogen Ralstonia solanacearum";  
RL Nature 415:497-502(2002).  
DR EMBL; AL646072; CAD16564.1; -;  
DR InterPro; IPR000873; AMP-bind.  
DR Pfam; PF00501; AMP-binding; 1.  
DR PROSITE; PS00455; AMP-BINDING; 1.  
KW Ligase; Complete proteome.  
SQ SEQUENCE 569 AA; 62332 MW; F9D8556300638F54 CRC64;

Query Match 62.1%; Score 41; DB 16; Length 569;  
Best Local Similarity 61.5%; Pred. No. 34;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CATLRVYKGGGXA 13  
| : : : : :  
Db 334 CSRLRVANGGMA 346

## RESULT 3

Q8D4T7 PRELIMINARY; PRT; 276 AA.  
ID Q8D4T7  
AC Q8D4T7  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Uncharacterized protein conserved in bacteria.  
GN W21202.  
OS Vibrio vulnificus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=672;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CMCP6;  
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,  
Choy H.E.;  
RT "Complete genome sequence of Vibrio vulnificus CMCP6";  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE016812; AA08099.1; -;  
KW Complete proteome.  
SQ SEQUENCE 276 AA; 31400 MW; E4A8CB7B4A6916 CRC64;

Query Match 60.6%; Score 40; DB 16; Length 276;  
Best Local Similarity 77.8%; Pred. No. 24;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 TLRVYKGGG 11  
| : : : : :  
Db 193 TTRVYRGGG 201

## RESULT 4

Q8DESO PRELIMINARY; PRT; 381 AA.  
ID Q8DESO  
AC Q8DESO  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Conserved hypothetical protein.  
GN VV10515.  
OS Vibrio vulnificus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=672;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CMCP6;  
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,  
Choy H.E.;  
RT "Complete genome sequence of Vibrio vulnificus CMCP6";  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE016798; AA09033.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 381 AA; 43301 MW; BC955683E9FB2FD CRC64;

Query Match 60.6%; Score 40; DB 16; Length 381;  
Best Local Similarity 77.8%; Pred. No. 34;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 TLRVYKGGG 11  
| : : : : :  
Db 295 TARVYRGGG 303

## RESULT 5

Q8Z8V6 PRELIMINARY; PRT; 384 AA.  
ID Q8Z8V6  
AC Q8Z8V6  
DT 01-MAR-2002 (TRENBLrel. 20, Created)  
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)  
DE Hypothetical protein SY0480.  
GN SY0480.  
OS Salmonella typhi.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OX NCBI\_TaxID=601;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CT18;  
RX MEDLINE=21534947; PubMed=11677608;  
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,  
Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
Whitehead S., Barrall B.G.;  
RT "Complete genome sequence of a multiple drug resistant Salmonella  
Enterica serovar Typhi CT18";  
RL Nature 413:848-852(2001).  
DR EMBL; AL627266; CAD08897.1; -;  
DR InterPro; IPR006597; Sel\_like.  
DR SMART; SM00671; Sel1; 4.

```
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 384 AA; 42954 MW; FE58E5A584CE6894 CRC64;

Query Match          60.6%; Score 40; DB 16; Length 384;
Best Local Similarity 72.7%; Pred. No. 34;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 TLRVYKGGGXA 13
   ||| ||| |||
DB 54 TLRQYAGGSA 64

RESULT 6
Q92UR7 PRELIMINARY; PRT; 607 AA.
AC Q92UR7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein RB1025.
GN RB1025 OR SWB21585.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoeelter F.J., Hernandez-Lucas I., Becker A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AL603645; CAC49425.1;
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 607 AA; 67013 MW; 88AFA2B6B00B8032 CRC64;

Query Match          60.6%; Score 40; DB 16; Length 607;
Best Local Similarity 87.5%; Pred. No. 56;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 LRVYKGGG 11
   ||| ||| |||
DB 518 LRVYRGGG 525

RESULT 7
Q9KKS7 PRELIMINARY; PRT; 619 AA.
AC Q9KKS7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical protein VCA1023.
GN VCA1023
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eelsen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;

"DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL; AE004428; AAF96919.1;
DR TIGR; VCA1023;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 619 AA; 71311 MW; DF5E5C1A07D4D013 CRC64;

Query Match          60.6%; Score 40; DB 16; Length 619;
Best Local Similarity 77.8%; Pred. No. 57;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 TLRVYKGGG 11
   ||| ||| |||
DB 536 TTRVYRGGG 544

RESULT 8
Q9LXL3 PRELIMINARY; PRT; 932 AA.
ID Q9LXL3
AC Q9LXL3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Kinesin-like protein.
GN F7K15_60.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Obermaler B., Ottenwaelder B., Duchemin D., Zeitler K., Meves H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL353871; CAB89042.1;
DR HSSP; P17119; 3KAR.
DR InterPro; IPR001752; kinesin_motor.
DR Pfam; PF00225; kinesin; 1.
DR PRINTS; PR00380; KINESINHEAVY.
DR SMART; SM00129; KISC; 1.
DR PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
DR PROSITE; PS00667; KINESIN_MOTOR_DOMAIN2; 1.
KW ATP-binding; Coiled coil; Microtubules; Motor protein.
SQ SEQUENCE 932 AA; 105805 MW; 721EC6C75FD762D4 CRC64;

Query Match          60.6%; Score 40; DB 10; Length 932;
Best Local Similarity 70.0%; Pred. No. 88;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 CATLRVYKGG 10
   ||| ||| |||
DB 83 CATQEVYEGG 92

RESULT 9
Q8LGU3 PRELIMINARY; PRT; 937 AA.
ID Q8LGU3
AC Q8LGU3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Kinesin-like protein.
GN TES.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
```

OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Ws2;  
 RA Yang C.Y., Spielman M., Coles J.P., Li Y., Ghelani S., Bourdon V.,  
 RA Brown R.C., Lemmon B.E., Scott R.J., Dickinson H.G.;  
 RA "TETRASPORE encodes a kinesin required for male meiotic cytokinesis in  
 RT Arabidopsis.";  
 RT Arabidopsis.";  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Ws2;  
 RA Yang C.Y., Spielman M., Coles J.P., Li Y., Ghelani S., Bourdon V.,  
 RA Brown R.C., Lemmon B.E., Scott R.J., Dickinson H.G.;  
 RA "TETRASPORE encodes a kinesin required for male meiotic cytokinesis in  
 RT Arabidopsis.";  
 RT Arabidopsis.";  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ496182; CAD42658.1; -;  
 DR EMBL; AJ507734; CAD45643.1; -;  
 DR InterPro; IPR001752; kinesin\_motor.  
 DR Pfam; PF00225; kinesin; 1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
 DR PROSITE; PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.  
 KW ATP-binding; Coiled coil; Microtubules; Motor protein.  
 SQ SEQUENCE 937 AA; 106428 MW; 827E0C7AB9FA819C CRC64;  
 Query Match 60.6%; Score 40; DB 10; Length 937;  
 Best Local Similarity 70.0%; Pred. No. 88;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CATLRYKGG 10  
 DB 83 CATQEVYEGG 92  
 RESULT 10  
 QBL5J2  
 ID QBL5J2 PRELIMINARY; PRT; 937 AA.  
 AC QBL5J2;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Kinesin-like protein.  
 GN TES.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Col-3; TISSUE=Flower;  
 RA Yang C.Y., Spielman M., Coles J.P., Li Y., Ghelani S., Bourdon V.,  
 RA Brown R.C., Lemmon B.E., Scott R.J., Dickinson H.G.;  
 RA "TETRASPORE encodes a kinesin required for male meiotic cytokinesis in  
 RT Arabidopsis.";  
 RT Arabidopsis.";  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Col-3;  
 RA Yang C.Y., Spielman M., Coles J.P., Li Y., Ghelani S., Bourdon V.,  
 RA Brown R.C., Lemmon B.E., Scott R.J., Dickinson H.G.;  
 RA "TETRASPORE encodes a kinesin required for male meiotic cytokinesis in  
 RT Arabidopsis.";  
 RT Arabidopsis.";  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ495781; CAD42234.1; -;  
 DR EMBL; AJ508243; CAD48111.1; -;  
 DR InterPro; IPR001752; kinesin\_motor.  
 DR Pfam; PF00225; kinesin; 1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00129; KISC; 1.

DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
 DR PROSITE; PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.  
 KW ATP-binding; Coiled coil; Microtubules; Motor protein.  
 SQ SEQUENCE 937 AA; 106389 MW; A251ACA2EB7F82E7 CRC64;  
 Query Match 60.6%; Score 40; DB 10; Length 937;  
 Best Local Similarity 70.0%; Pred. No. 88;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CATLRYKGG 10  
 DB 83 CATQEVYEGG 92  
 RESULT 11  
 QBLN22  
 ID QBLN22 PRELIMINARY; PRT; 938 AA.  
 AC QBLN22;  
 DT 01-OCT-2002 (TREMBlrel. 22, Created)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Kinesin-like protein.  
 GN ATNACK2.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21952493; PubMed=11955449;  
 RA Nishihama R., Soyano T., Ishikawa M., Araki S., Tanaka H., Asada T.,  
 RA Irie K., Ito M., Terada M., Banno H., Yamazaki Y., Machida Y.;  
 RT Expansion of the cell plate in plant cytokinesis requires a kinesin-  
 RT like protein/MAPKK complex.";  
 RL Cell 109:87-99(2002).  
 DR EMBL; AB088121; BAC03248.1; -;  
 DR InterPro; IPR001752; kinesin\_motor.  
 DR Pfam; PF00225; kinesin; 1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
 DR PROSITE; PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.  
 KW ATP-binding; Coiled coil; Microtubules; Motor protein.  
 SQ SEQUENCE 938 AA; 106520 MW; 8C570771E1040C78 CRC64;  
 Query Match 60.6%; Score 40; DB 10; Length 938;  
 Best Local Similarity 70.0%; Pred. No. 88;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 CATLRYKGG 10  
 DB 84 CATQEVYEGG 93  
 RESULT 12  
 QBIU03  
 ID QBIU03 PRELIMINARY; PRT; 173 AA.  
 AC QBIU03;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE Hypothetical protein Y73B6BL.9b.  
 GN Y73B6BL.9.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;

RA Waterston R.;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]

RP SEQUENCE FROM N.A.  
RC STRAIN-Bristol N2;  
RA Graves T.;  
RT "The sequence of C. elegans cosmid Y73B6BL.";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBSJ databases.  
RN [3]

RP SEQUENCE FROM N.A.  
RC STRAIN-Bristol N2;  
RA Waterston R.;  
RT Submitted (NOV-2002) to the EMBL/GenBank/DBSJ databases.  
RL EMBL: AC084197; RAN63427.1;  
DR Hypothetical protein.  
KW SEQUENCE 173 AA; 18222 MW; 0F73163B034C59E2 CRC64;

SQ SEQUENCE 173 AA; 18222 MW; 0F73163B034C59E2 CRC64;  
Query Match 59.1%; Score 39; DB 5; Length 173;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 RVYKGGG 11  
DB 28 RVYKGGG 34

## RESULT 13

Q8MXR7 PRELIMINARY; PRT; 241 AA.

AC Q8MXR7;  
DT 01-OCT-2002 (TREMBLrel. 22, Created)  
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Hypothetical protein Y73B6BL.9.

GN Y73B6BL.9.

OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI\_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Bristol N2;

RX MEDLINE-99069613; PubMed-9851916;

RA Waterston R.;

RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-Bristol N2;

RA Graves T.;

RT "The sequence of C. elegans cosmid Y73B6BL.";

RL Submitted (OCT-2000) to the EMBL/GenBank/DBSJ databases.  
RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-Bristol N2;

RA Waterston R.;

RL Submitted (MAY-2002) to the EMBL/GenBank/DBSJ databases.

DR EMBL: AC084197; AAM44399.1;

DR InterPro: IPR005818; Histone\_H1/H5.

DR InterPro: IPR005819; Histone\_H5.

DR InterPro: IPR003216; Linkerhist\_N.

DR PRINTS: PR00624; HISTONEH5.

DR ProDom: PD000373; Linkerhist\_N; 1.

DR SMART: SM00526; H15; 1.

KW Hypothetical protein.

SQ SEQUENCE 241 AA; 25409 MW; 5E738DFDEE1A692 CRC64;

Query Match 59.1%; Score 39; DB 5; Length 241;  
Best Local Similarity 100.0%; Pred. No. 32;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 RVYKGGG 11  
DB 28 RVYKGGG 34

## RESULT 14

Q98EM3 PRELIMINARY; PRT; 285 AA.

AC Q98EM3;  
DT 01-OCT-2001 (TREMBLrel. 18, Created)  
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Acetate CoA-transferase, alpha subunit.  
GN MLI4183.

OS Rhizobium loti (Mesorhizobium loti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Phyllobacteriaceae; Mesorhizobium.

OX NCBI\_TaxID=381;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-WAFF303099;

RX MEDLINE-21082930; PubMed-11214968;

RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,

RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,

RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matuno A.,

RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,

RA Takeuchi C., Yamada M., Tabata S.;

RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium

RL Mesorhizobium loti.";

RL DNA Res. 7:331-338(2000).

DR EMBL: AP003003; BAB50895.1;

DR InterPro: IPR004165; CoA\_trans.

DR Pfam: PF01144; CoA\_trans; 1.

KW Transferase; Complete proteome.

SQ SEQUENCE 285 AA; 30987 MW; 2EE611B692106EED CRC64;

Query Match 59.1%; Score 39; DB 16; Length 285;

Best Local Similarity 53.8%; Pred. No. 39;

Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CATLRVYKGGGXA 13

DB 120 CAVERGYRGAGLA 132

## RESULT 15

Q8TOY4 PRELIMINARY; PRT; 456 AA.

AC Q8TOY4;

DT 01-JUN-2002 (TREMBLrel. 21, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Dopamine receptor type D2.

GN DOP2.

OS Apis mellifera (Honeybee).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Hymenoptera; Apoidea; Apoidea;

OC Apidae; Apis.

OX NCBI\_TaxID=7460;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain mushroom bodies;

RX MEDLINE-98194783; PubMed-9535160;

RA Ebert P.R., Rowland J.E., Toma D.P.;

RT "Isolation of seven unique biogenic amine receptor clones from the

RT honey bee by library scanning.";

RL Insect Mol. Biol. 7:151-162(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain mushroom bodies;

RA Humphries M.A., Mustard J., Hunter S.J., Mercer A., Ward V.;

RA Ebert P.R.;  
RT "An invertebrate D2 type dopamine receptor exhibits plasticity of  
RT expression in the mushroom bodies of the honey bee brain correlated  
RT with behavioural maturation of adults."  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
DR EMBL; AF498306; AAM19330.1; -  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm\_1; 1.  
DR PROSITE; PS00237; G\_PROTEIN\_RECEP\_F1\_1; 1.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW G-protein coupled receptor; Receptor; Transmembrane.  
SQ SEQUENCE 456 AA; 52309 MW; 44E6D35470B367CA CRC64;

Query Match 59.1%; Score 39; DB 5; Length 456;  
Best Local Similarity 66.7%; Pred No. 63;  
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 TLRYVYKGG 11  
||||:||||  
Db 257 TLRIHRRGG 265

Search completed: August 28, 2003, 18:38:02  
Job time : 33.3182 secs